We thank Edward Zaron for his very detailed and constructive review, with in-depth comments and suggestions very helpful to propose a second version of the manuscript. We propose here a point-by-point response, supported with new tests, computations and figures when necessary.

- *No details seem to be provided about the Q matrices for both the mesoscale and tidal signals. Can you provide, say, a pseudo-spectrum of variance that shows how Q varies as a function of wavenumber or horizontal mode number; or do you feel that this is represented equivalently with the physical-space representers shown in Fig 5? Can you provide maps of the mesoscale and tidal variance (i.e., the diagonal elements of the Q matrix)? Did you perform any tuning to adjust the ratio of Q and R, or the ratio of the tidal Q and the mesoscale Q?*

Indeed, the discussion about the Q matrix was not detailed enough in the manuscript, only mentioning that the diagonal terms of Q are designed to approach equivalent covariances of Ocean mesoscale altimetry signal. We propose a more detailed description, which is based on an input dataset of power spectral densities measured by the AltiKa along-track SSH anomalies. For a given dominant wavelength of the wavelet of the mesoscale decomposition, the corresponding column of the Q matrix has the energy of the integrated spectrum between half the preceding wavelength and half the following wavelength. The database of power spectral densities is illustrated on Figure 1 (in space for a given wavelength) and Figure 2 (in wavelength at a given location) of this response document.

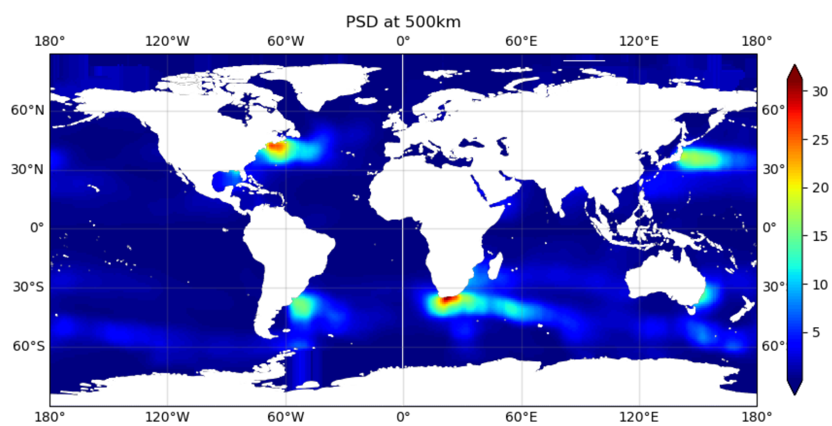


Figure 1 : power spectral density (in cm²/cy/km) at 500km wavelength from the along-track sea level anomalies of AltiKa satellite

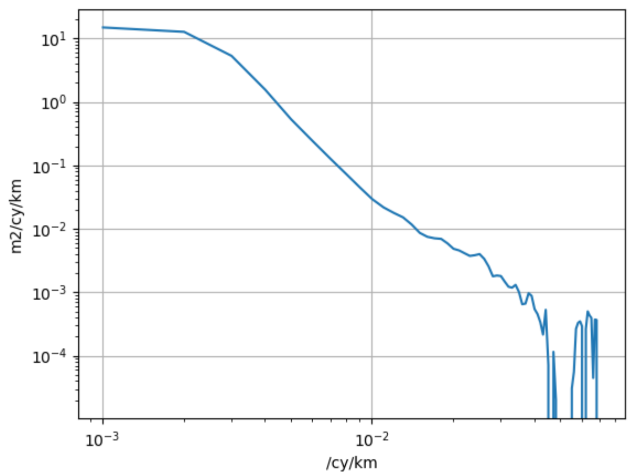


Figure 2 : power spectral density (in cm²/cy/km) at 60°W, 35°N from the along-track sea level anomalies of AltiKa satellite

As shown on the figures, there is an important spatial variability of the energy, and an energy decrease with the wavenumber (red spectrum). Note that the Altimeter noise was removed following Fu and Xu, 2012. There variabilities are therefore taken into account in Q, which allow a representer function (Figure 4 of the manuscript).

As for internal tides, in the Q matrix block, we did not consider any a priori on regional variability. We set up the diagonals of Q at 2cm squared for mode 1, 1cm squared for mode 2, for all tidal components, as we propose to explain in the new version of the manuscript.

If I understand correctly, the estimate you compute using eqn (8) or (13) is biased towards zero. A symptom of this bias is the observation that the explained variance is larger than the signal variance (as you noted with regard to the bottom panel of Fig 8). Can you plot a map of this variance ratio and interpret it with regard to either the bias or the tuning of Q? Would you consider using a non-zero estimate of the tide or the mesoscale in order to reduce this bias?

Yes, this is a good point, we plotted below (Figure 3) the variance of the signal (top panel) and the difference between this latter and the variance reduction after applying the correction (we found that plotting the difference was more explicit than the ratio, but the message is the same) : indeed, the bottom panel has more blue color than red, meaning that the variance reduction is stronger than the variance of the signal itself. However, some clear red spots are present, in particular in the Kurushio. One hypothesis is that the mesoscale leakage (although mitigated) gets sometimes higher that the bias toward zero.

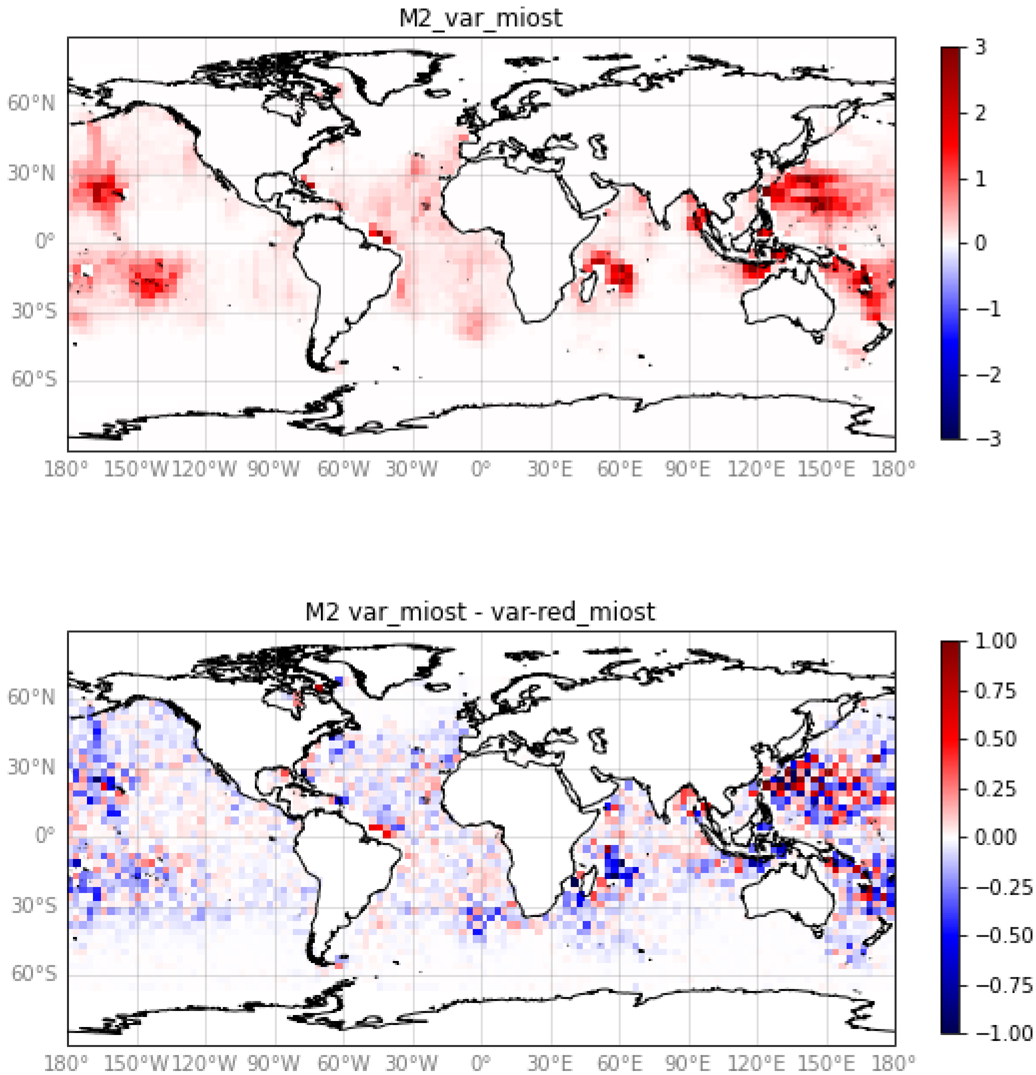


Figure 3 : Signal variance (cm2) of the M2 estimation binned on a 2° by 2° grid (top panel) and difference between the signal variance and the reduction of variance after applying the signal to the independent data (lower panel). The blue zones indicate a higher variance reduction than the variance of the signal itself.**

By curiosity, we did the same figure, but with the HRET model (Figure 4). We still note the presence of red spots, and some of the blue spots (variance reduction higher than the signal variance) are more blue than in MIOST, meaning that the bias feature is present in the two models.

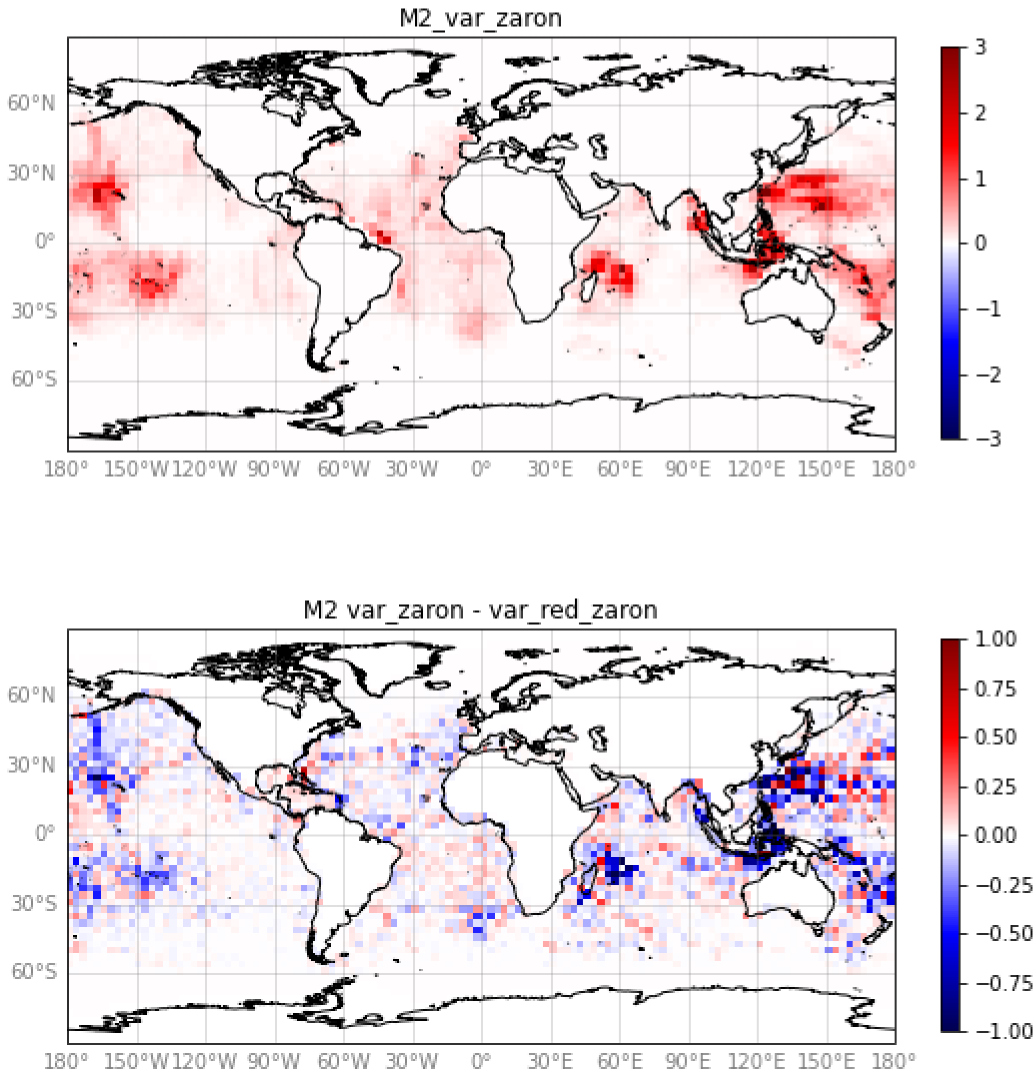
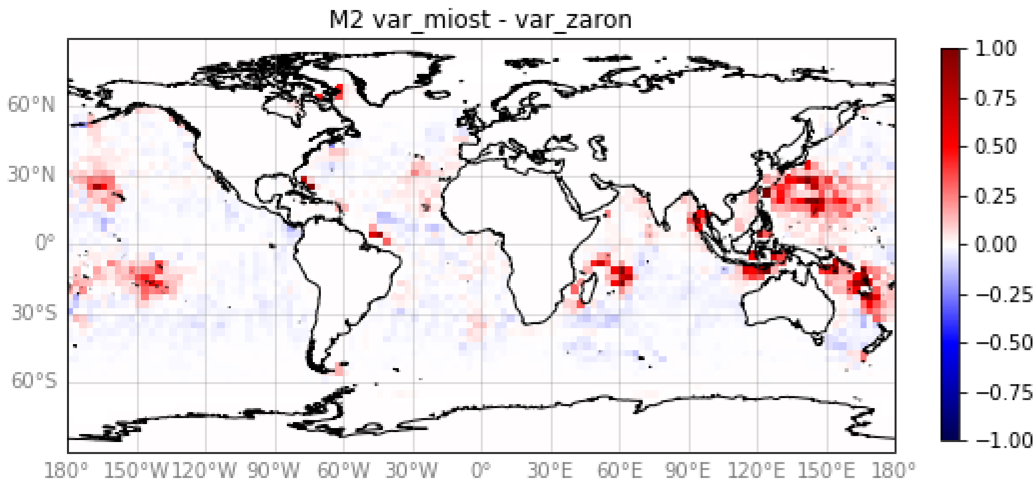


Figure 4 : Same as Figgure 3, for the HRET (Zaron) M2 signal.

It is also interesting to plot the difference of signal variance between MIOST and HRET (lower panel of Figure 5) compared with the difference of variance reduction. We can note that the signal in MIOST is significantly stronger (much more red than blue in the upper panel) but the variance reduction is only slightly stronger (lower panel). This would suggest that MIOST brings overall more signal, but also more error with a light net advantage in variance reductions. This also supports that the two models have probably their own advantages.

As for using a non-zero prior estimate of the tides or mesoscale, we are not sure this would mitigate the bias issue since those estimate have to be generated with Altimetry observation containing both signals anyways, the the bias issues would be still there. We would be happy to discuss more this question is this is not clear.



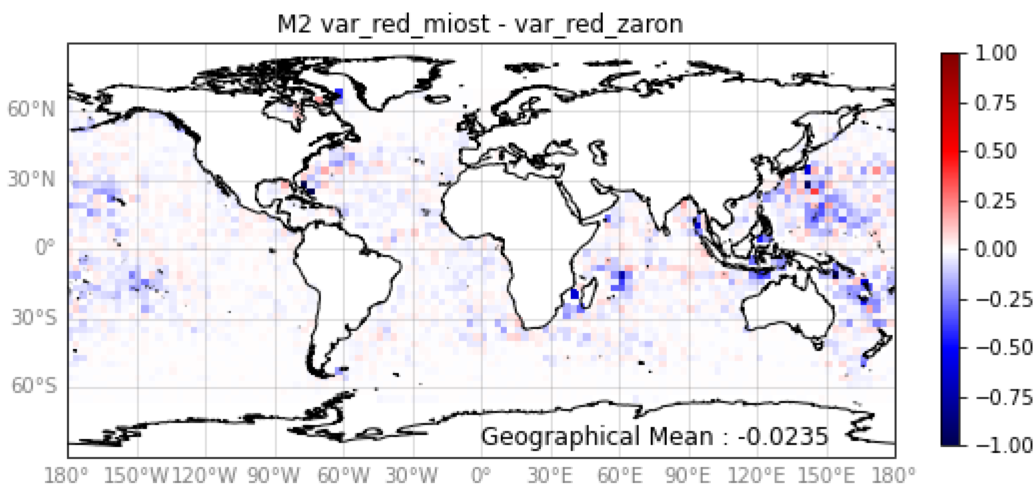


Figure 5 : difference of signal variance between HRET and MIOST (top). The red zones indicate where the MIOST signal has more variance. Difference of variance reduction (as the bottom-left panel of Figure 10 in the paper)

No details are provided with regard to the time-dependence of the tides, except for equation (4). It appears that the nodal modulations have been omitted, but this is a substantial effect over a 25 year record. Properly accounting for this would probably further increase the explained *variance of the tidal estimates with respect to both the validation data and the assimilated data; and it should furthermore reduce the (low) bias of the tidal amplitudes.*

This is correct, and for Eq.4, we rigorously assume stationarity in the tidal forcing and solution we seek as this is an idealized test.
However, in the main part of the paper for the Altimetry implementation, the nodal modulations have been accounted, with an amplitude and phase slightly varying. The nodal variations have been implemented with the pytide module (https://github.com/CNES/pangeo-pytide). We propose to clarify this in the manuscript, thanks for this good point!
However, we did not note a sustancial effect with respect to no variations (we did the test on one tile), only a minor effect with less than a few percent differences in the performances.

The English language usage is sometimes awkward or non-standard, especially with regard to capitalization. I am not evaluating it or going to list all the potential edits during this reading.

Thanks for noting, we hope we fixed all capitalization issues, and tried our best to improve the English language in the new version.

Smaller comments:

l14-15: Not sure where they get the 70% phase-modulated.

Indeed this number does not appear in the 2017 Zaron paper. We propose to comment the Figure 9 from this same paper, suggesting that the fraction of unstationary internal tides exceeds 50% in Equatorial regions and Strong Western boundary currents. 70% for a global average was certainly an overestimation from a mis-reading of that figure.
This is clarified in the new manuscript.

l35: "covariances" --> "spatial covariances"?
Correct

l49: specify, "x_i and x_j are uncorrelated for i \ne j"
Yes, and we also mentioned “two components” instead of “N”. This is also corrected.

l62: Here it is specified that x_1 refers to the time-series of a scalar. Aha. But after line 70, it is clear that x_2 is a two-component vector containing the harmonic constants of the high-frequency component.
Yes, and see our answer above : it was indeed unclear as we mentioned two components earlier.

l75: The reference to localization should either be dropped or explained precisely what is meant.

Yes indeed, we initially wanted to point out the necessity of inverting globally, as opposed to many situations where localization is implemented for practical reasons, But this is already discussed later. We agree it is more clear not to mention the non-localization here.

p4, last line: How does it differ from harmonic analysis? If it is identical, then say so.
Not identical. We propose the following explanation:
For the harmonic part, this separate estimation \added{performs nearly as a harmonic analysis, except that the finite $\mathbf{B}$ and non-zero $\mathbf{R}$ matrices tend to slightly reduce noise contamination with respect to harmonic fit.}

l80: By "sequential estimation" do you mean that the low-frequency component is estimated by itself from the entire time series, and then this estimate is subtracted before estimating the high-frequency component? This usage of "sequential" is confusing since the term might also refer to sequential estimation (i.e., a Kalman filter) which sequentially processes the observations in time.
Yes. Indeed, we did not realize the potential confusion. We could propose “alternate estimation” which is actually closer to what is done, since the sequence is done in alternate order for each component evaluation.

l95-l100: This is very good discussion of bias in this context.
Thanks!

l125: Can you support your assumption that no correlation exists between the components (l116) by saying that the \Gamma_k are chosen to approximately diagonalize the state covariance? If you could provide some observational data to support the choice of \Gamma_k, that would be even better!
Aha: now I see the mention of this later, around l130.

The different justifications were indeed incomplete and appeared in misleading order. We propose to re-write the whole paragraph presenting the Q matrix, directly mentioning the two main assumption : (1) the components are assumed orthogonal and (2) each given component is diagonalized (by construction). This is why the global Q matrix can be diagonal.

l161: I believe the reference should be to Fig 4, not Fig 3? You will probably need a reference or short discussion to explain what is a "representer".

Correct. We propose to add a description of the representor, but in the previous paragraph (regarding mesoscales) when it is introduced for the first time. In this paragraph, we also propose some clarifications, as the representer is shown for a given baroclinic mode (the first one) and for a given tidal component (M2). We believe the whole paragraph needed some clarifications, as proposed in the new version.

Fig 4: Was the representer shown in the right panels constructed from the 12-equiangular basis elements? I am surprised that it is as radially-symmetric as shown.

Yes, we believe that the number of directions must be related to the spatial extension of the plane waves. The longer it is, the higher the number of directions is needed to fulfill the decomposition. We propose a clarification of this point in the new manuscript.

l173: This is a good compromise between domain size and degrees of freedom.
Yes, we initially tried to inverse over 30 by 30 degree tiles, but the computation was too heavy...

Fig 5: Did you subsample or average the observations in the along-track direction? Or did you use 1 Hz data? Why not show the same lat/lon window in each panel?

Oh yes, we did, and actually a paragraph was clearly missing regarding the input dataset, as pointed out by another reviewer. As for the subsampling, we considered ⅓ Hz after averaging every three points. We verified that this averaging did not impact the results (sensitivity tests, not shown, suggested an impact beyond 4 consecutive points averaged). The new version of the manuscript has a new paragraph presenting the dataset.

l231: How were the diagnonals of the Q matrices chosen initially? What information was used to estimate the variances of the signals?

They were chosen according the the along-track Altimetry spectral database. Now that we propose a paragraph of clarification for the Q matrix, we believe this point l231 is also clarified.

Fig 8 and discussion: Usually "signal variance" refers to the data, but I believe you are using it to refer to the variance of the estimated signal. Perhaps this could be clarified. My interpretation is based on the fact that you note the explained variance is larger than the "signal variance" in cases 2 and 3.
Yes, we propose to use the following terminologies : “variance of the solution” and “explained variance after applying the solution as a correction to independent altimetry data”. We hope now the figure and referring text is more clear.

Maybe I missed it, but no where do I see discussion of what altimeter missions were used. It looks like CryoSat-2 is in the post-2017 validation dataset, but is this the only mission used?

Yes, we realized that the validation dataset was not properly described. New sentences have been added, and also refer to the new paragraph describing the main dataset used to compute the solution. The dataset refers to all satellite missions post September 2017.

l235: Why are you using only a year for the validation period?
It was a typo, it is actually 2 years. We performed the analysis in 2020 and at that time, our series ended in December 2019.

Fig 10: Please state the units of the comparisons (cm^2, I think?).

Yes, done!

Why does the bias problem not seem to be as large as suggested by Fig 5? Perhaps I do not understand your sequential estimates, and they differ more significantly from the approach used in HRET. Or, maybe your low-frequency solution obtained here is quite different from the Duacs/Ssalto-based mesoscale correction used in HRET. Based on your Fig 5, I would have expected your estimate to explain a lot more variance than HRET.

We think this point is now better supported by the new analysis, especially the Figure 4 showing that the MIOST solution is globally more energetic than HRET, especially in zones of high internal tides. So even id the net gain of variance reduction is overall moderate, the signal is stronger. In other words, we could say that MIOST resolves overall more signal, but also introduces slighly more errors.

Table 1: What do the percentages refer to (is the decimal point placed correctly?)? Please label the sub-tables with M2 and K1.
We propose a new table in the next version of the manuscript, more synthetic and with correct number of digits for percentage display.